Defense Technical Information Center Compilation Part Notice

ADP010538

TITLE: Image Based Contrast-To-Clutter Modeling of Detection

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Search and Target Acquisition

To order the complete compilation report, use: ADA388367

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, ect. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP010531 thru ADP010556

IMAGE BASED CONTRAST-TO-CLUTTER MODELING OF DETECTION

David L. Wilson
US Army CECOM NVESD
10221 Burbeck Rd., STE 430
Ft. Belvoir, VA 22060-5806
Phone: (+) 703 704 2106
Email: dwilson@nvl.army.mil

1. SUMMARY

Using image-based metrics, contrast-to-clutter modeling is applied to the Search-2 visible image set and perception experiment data. To calculate the contrast metric, a new image is generated from the original image by replacing the target with an "expected background" using the local background surrounding the target and the natural horizontal correlation present in most surface-to-surface scenes. The contrast metric is obtained from the difference of this new image and the original image. Via a simple mathematical formula, the ratio of the contrast measure to a clutter metric is used to predict performance.

Keywords: clutter, line PSS, contrast-to-clutter, power spectrum signature, PSS, probability of detection, RSS

2. INTRODUCTION

Models to predict probability of detection by human observers viewing images vary from contrast and resolution methodology, as in Johnson criteria for predicting IR FLIR performance, to complex visual models that attempt to consider the target and background spectral natures and the effects of those on the human detection process. Current modeling tends to have less than the desired accuracy in predicting probability of detection.

The work presented here originates from metrics used experimentally at NVESD for predicting IR FLIR performance in real imagery with cluttered backgrounds. The metrics used here were introduced at the 1997 SPIE Meeting (Ref. 1) and the 1998 Army Science Conference (Ref. 2). These metrics are combined together for the first time against a visible data set and perception data.

As the metrics and earlier modeling were against IR instead of visible imagery, the results here required the metrics to be evaluated against grayscale renditions of the original color images. It is possible that color equivalents for the metrics used might be found with further experimentation. However, for most of the images in the Search-2 data set, it is suspected, and somewhat confirmed by the following analysis, that color may not be a major factor in this image set.

A second caveat in this analysis is that the data set is relatively small with most targets having relatively high probability of detection. A broader range of probability of detection would be more desirable for a robust evaluation.

Finally, as the originally provided data set had errors in the ranges for images 35, 38 and 43, these images were excluded from the analysis as time did not permit using the corrected ranges that were later provided for those three images.

3. IMAGE METRICS

Image metrics used here fall into four categories: contrast metrics, clutter metrics, size metrics and shape metrics. The following will introduce the objective contrast, clutter and size metrics used in the modeling.

3.1. Contrast Metrics

In general terms, a contrast metric is a metric that measures the intensity difference between a target and its local background. Such metrics may be in gray levels, light intensity levels, or in the case of IR FLIR imagery in temperature units. Usually, the local background is taken to be a box with dimensions (width and height) the square root of 2 multiplied by the dimensions (maximum width height) of the target. In the case of a rectangular target, this gives the local background the same area as that of the target.

The simplest contrast metric is the difference between the target and background means:

$$\Delta\mu = \left|\mu_{tgt} - \mu_{bkg}\right| \qquad \text{(Eq .1)}$$

where μ_{tgt} is the target mean intensity and μ_{bkg} is the background mean intensity.

The difficulty with the above metric is that it does not consider internal structure of the target and background. The target and background may have the same means but the target may be detectable due to its internal structure. One of the most commonly used contrast measures that attempts to somewhat correct this problem is the RSS (Root Sum-of-Squares). The RSS is given by:

$$RSS = \left[\frac{1}{POT}\sum_{pixel(i,j)\in tgt}(t_{i,j} - \mu_{bkg})^2\right]^{1/2} \qquad \text{(Eq. 2)}$$

where $t_{i,j}$ is the intensity of the pixel (i,j) and POT is the number of pixels on target. The RSS can be calculated readily from the target and background means and the target standard deviation by the following alternative formula:

$$RSS = \left[\left(\mu_{tgt} - \mu_{bkg}\right)^2 + \sigma_{tgt}^2\right]^{1/2} \qquad \text{(Eq. 3)}$$

where σ_{tgt} indicates the standard deviation of the target.

A different contrast metric that has been proposed (Ref. 1) is an implementation of the PSS (Power Spectrum Signature). This requires us to define an "expected background" image for what we might expect to see if the target were not present.

Paper presented at the RTO SCI Workshop on "Search and Target Acquisition", held in Utrecht, The Netherlands, 21-23 June 1999, and published in RTO MP-45.

Naturally, such an image cannot be precisely exactly defined. For the moment, assume we have such an expected background image; we may then define the PSS as:

$$PSS = \left[\frac{1}{POT} \sum_{i,j} \left(t_{i,j} - b_{i,j} \right)^2 \right]^{1/2} \quad \text{(Eq. 4)}$$

where the summation is over all pixels which are different in the original image and expected background image, t_{pix} indicates the intensity of target pixels and b_{pix} indicates the intensity of expected background pixels.

The problem is to now define a usable concept of expected background. Note that this is not the same as the actual background. For example, the actual background might contain a hot or bright rock. We would not expect to see such. The expected background should be one that does not draw the attention of the observer. There are various possible implementations for an expected background. For example, one might replace the target by the mean of the local background. If one actually does this, one quickly discovers that the flat intensity that results is far from expected and readily draws one's attention.

The implementation of the expected background that is used here is based on the fact that real images of surface-to-surface scenes tend to have high horizontal correlation. This is probably due to two primary factors. The first is basic geometry. Even a solid circle patch on the ground at a distance horizontally will appear to be an ellipse with the major axis in the horizontal direction. Another factor to consider is that local to the target, the horizontal will tend to be at the same range and suspect to the same propagation effects as well as contain the same vegetation. There are exceptions to this: for example, long exposed tree trunks would give a strong vertical correlation in that part of an image. In general though, the horizontal correlation might be expected.

The above leads to the concept of the (horizontal) line expected background and PSS. We define the expected ground intensity at target pixel locations to be:

$$b_{i,j} = \left(1 - {i'}/{n(j)} \right) \mu_L + \frac{i'}{n(j)} \mu_R \quad \text{(Eq. 5)}$$

where i' is the distance along the horizontal from the left edge of the target to pixel (i, j), n(j) is the distance along the j^{th} horizontal inside the target, μ_L is the mean intensity on the horizontal in the local background to the left of the target and μ_R is the mean intensity on the horizontal in the local background to the right of the target (see Fig. 1). In concept, the line PSS is a linear horizontal interpolation between the mean local left intensity and the mean local right intensity.

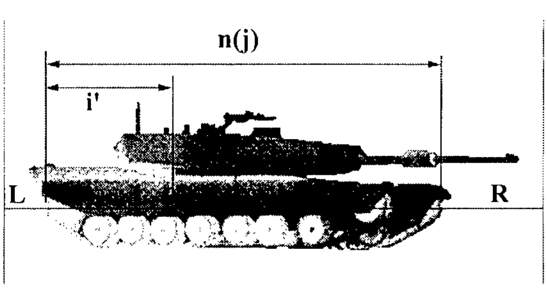

Figure 1: Calculating Line PSS

The image formed with the line PSS expected background often works surprisingly well. For example, see Fig. 2a, Fig. 2b, Fig. 3a and Fig. 3b.

Figure 2a: Crop of target in image 10

Figure 2b: Line PSS expected background of Fig. 2a

As expected, if the target is too close (large) and surrounded by clutter, the expected background often appears too flat or to have unusual horizontal streaks. Also, if there is a high contrast clutter object horizontally in the local background of the target, this will cause a conspicuous horizontal streak in the expected background image. In such cases, the line PSS expected background methodology needs to be modified; but this has not been done in the analysis that follows.

3.2. Clutter Metrics

The simplest clutter metric is the standard deviation of the image or the standard deviation of the local background σ_{bkg}. In practice, this does not seem to work very well.

Dr. Silk at the Institute for Defense Analysis has proposed another clutter metric (Ref. 3). It is a modification of the Schmieder-Weathersby clutter metric (Ref. 4) and bares some similarity to the form of the line PSS. As we will later form the ratio of the contrast metric to the clutter metric, similarity in form is a desirable property.

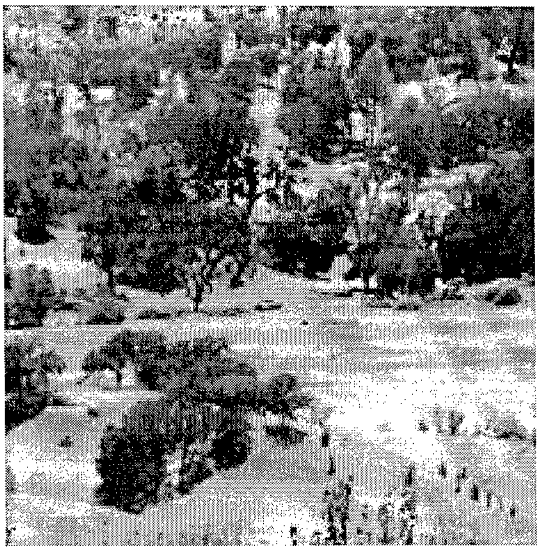

Figure 3a: Crop of target in image 33

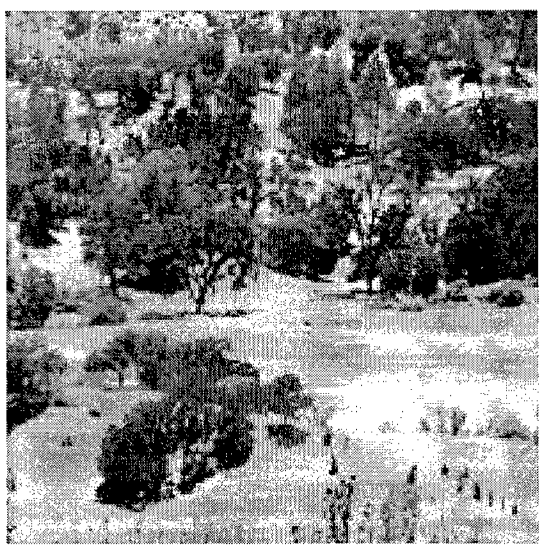

Figure 3b: Line PSS expected background of Fig. 3a

The implementation of this clutter metric may be visualized as a calculation obtained by convolving a square box centered on an image pixel through the image. There are two ways to handle difficulties encountered at the image border. The first is to pad the image by reflecting the image, or padding with the mean or some other constant. The second is to only convolve the box to positions that keep the box interior to the image. As little difference it obtained between these two choices and the second is easier to implement, it is the one used in the following analysis. The formula for the implemented clutter metric, denoted CM, may be expressed as:

$$CM = \left[\frac{\sum_i \sum_j \left(b_{i,j} - \overline{B}_{i,j} \right)^2}{N} \right]^{1/2} \qquad \text{(Eq. 6)}$$

where $b_{i,j}$ is the intensity of the (i, j) pixel, $B_{i,j}$ is the square box of pixels centered at (i, j) with the bar above it indicating the mean intensity of the box and N is the number of boxes convolved in the image.

As the target should not be considered cluttered, the clutter metric CM is actually calculated on the line PSS image having the expected background. This is even more important when multiple targets are present in the image, as not eliminating them from the calculation will cause too large a value for the clutter metric. Also, horizons should not be allowed in the calculation of CM, as they will dominate the calculation. One could similarly argue that other strong contrast, clearly recognized structures should also be eliminated from the CM calculation, as they probably do not act as clutter. But no procedure is currently proposed for doing this.

The size of the box has been subject to experimentation. Generally it has been set at 4 meters at the target range. This is larger than most targets. But should the box represent actual target size, expected target size or some other size? Clearly more work needs to be done concerning the box size; but in the past, using 4 meters at target range seems to have worked as well as any other choice and is used here.

Another issue is whether the clutter metric should be calculated over the entire image or over some region more local to the target. In the analysis that follows, CM was calculated over both the entire image and over a region subjectively local and similar to that near the target.

3.3. Target Size

The target size used in this analysis was the square root of the pixels on target. The software used to calculate the line PSS and CM required using a re-sized to 768x512 gray scale version of the images and re-segmentation of the targets. The square-root number of pixels on target (POT) from that segmentation was used as the target size.

3.4. Metric Values

Although the next section will model probability of detection using metrics described in this section, for completeness, a table of measured metrics is included here.

In Table 1, CM1 indicates the clutter metric calculated on the entire image while CM2 indicates the clutter metric calculated on a subjectively determined region containing the target that is typical to the area around the target. As one readily sees, CM is sensitive to the region chosen and that has caused some concern. But in truth, the difficulty is a lack of knowledge on what should really be considered clutter, rather than a problem with the calculation.

As mentioned earlier, images 35, 38 and 43 are excluded due to errors in the originally provided ranges that are used in determining the box sizes in the calculation of CM. Both the contrast and clutter metrics are measured in gray levels.

4. PROBABILITY OF DETECTION MODELING

Various combinations of the metrics might be considered for modeling probability of detection. Among these are contrast (such as RSS or line PSS) alone, size (such as square root of POT) alone, contrast times size and contrast times size divided by a clutter metric (such as the standard deviation of the local background or CM). This last predictor for probability of detection is loosely referred to as contrast-to-clutter. Of particular interest is the case of line PSS as the contrast metric and CM as the clutter metric. But other predictors will be also considered for comparison.

In each case, one generally expects the larger the above predictor, the larger the probability of detection. The natural goal is to find one that works "best" in general. "Best" can be defined in various, often conflicting, ways: least scatter, least

Table 1: Measured Metrics

Image	POT	RSS	Line PSS	CM1	CM2
1	38	38.95	32.76	6.8	8.53
2	60	32.99	34.66	5.88	8.13
3	42	43.71	41.42	6.17	7.58
4	14	19.07	29.58	7.28	9.66
5	707	60.82	68.54	17.04	21.43
6	24	23.03	30.09	10.64	12.85
7	13	17.7	24.95	7.44	9.02
8	22	30.7	32.98	8.29	9.9
9	899	35.33	40.31	14.66	16.59
10	79	43.74	48.24	9.37	12.13
11	43	21.79	26.52	7.77	10.79
12	452	44.91	61.04	16.81	19.47
13	152	25.36	36.13	10.57	11.04
14	185	35.34	38.74	11.51	14.72
15	16	21.17	25.07	5.54	6.75
16	77	33.57	39.79	7.53	9.32
17	80	43.86	52.47	12.89	11.81
18	193	28.29	51.13	7.88	15.56
19	146	20.67	19.45	13.84	14.65
20	203	50.92	58.62	12.48	13.16
21	15	41.56	44.98	8.54	10.31
22	18	15.69	29.65	7.52	8.3
23	14	30.05	47.25	9.07	11.19
24	25	27.19	25.73	10.68	6.96
25	37	38.93	49.34	10.3	10.14
26	37	24.62	22.37	8.03	9.76
27	19	36.38	49.51	6.93	8.7
28	80	36.96	29.84	11.75	13.38
29	17	36.67	37.64	8.74	9.91
30	34	30.4	44.38	9.9	12.76
31	245	39.84	39.78	12.79	20.42
32	7	32.89	35.27	8.75	10.33
33	47	48.96	44.8	11.24	14.41
34	1162	70.83	87.93	23.66	25.95
36	92	35.07	59.05	12.91	17.13
37	438	48.47	79.18	17.46	30.21
39	22	14.77	24.37	7.88	13.67
40	144	63.92	65.39	14.49	12.16
41	388	38.34	40.84	17.64	23.55
42	42	38.16	48.98	11.33	14.72
44	20	37.24	32.08	7.11	8.81

RMS error or largest correlation between measured and predicted probabilities.

For the modeling that follows, the equation

$$PD_{pred} = \frac{(X/X_{50})^E}{1+(X/X_{50})^E} \quad \text{(Eq. 7)}$$

is used to predict the probability of detection. In the equation, X represents the predictor or combination of metrics used in the prediction. Both E and X_{50} are determined by non-linear regression (the Levenberg-Marquadt least-squares method).

In the table below, r is the Pearson product-moment correlation, COD is the coefficient of determination (r^2) and r_s is the Spearman rank correlation. The larger the value of r, the better the correlation between the predicted and measured probabilities.

Table 2: Prediction Models

X	X_{50}	E	COD	r	r_s
RSS	17.2	3.63	0.65	0.81	0.58
RSS•√POT	71.2	2.12	0.71	0.84	0.77
RSS•√POT/σ_{bkg}	3.20	1.58	0.55	0.74	0.49
PSS	20.4	3.20	0.42	0.64	0.59
PSS•√POT	93.5	2.37	0.62	0.79	0.77
PSS•√POT/σ_{bkg}	4.45	1.70	0.52	0.72	0.58
PSS•√POT/CM1	9.76	1.91	0.42	0.65	0.65
PSS•√POT/CM2	9.68	2.72	0.63	0.79	0.70

Note that COD measures the relative amount of the measured variance accounted for by the model only if one assumes the residuals follow a normal distribution with constant variance. The Spearman rank correlation has no such requirement on the distribution and variance.

It is interesting that all the better models above may be working roughly equally well when one considers the sample size. Some of the models yielding the larger correlations above do not include a clutter metric. One would conjecture they would perhaps not do as well if there were a greater range in image clutter. Additionally, if the "gain" of the display is adjusted, the contrast-to-clutter metrics have the advantage that this effect should somewhat cancel. Within some reasonable range, assuming relative linearity between gray levels and screen brightness this seems reasonable.

Using CM2 instead of CM1 appeared to give a better correlation. Recall that CM2 is the clutter metric CM measured over a subjectively determined region containing the target that was judged similar to the area near the target, rather than using the entire image as in CM1. This region contained a tree line if the target was near one; otherwise, it did not. If there were trees near the target, the region contained as large a group of similar trees as could be selected. Different people might select different regions. One might hope to formalize this process. As was mentioned earlier, a better understanding, and probably also of measurement, of clutter is needed.

Figures 4a-b shows plots of PSS•√POT/CM2 versus measured probability of detection for the perception experiment data. The curve is the model prediction regression line. Figure 5a-b are similar plots of the case of RSS•√POT for comparison.

Although both the RSS•√POT based and PSS•√POT/CM2 based models appear to work nearly equally well, as noted earlier, the contrast-to-clutter model should in theory have advantages when there are large variations in clutter or changes in the display gain.

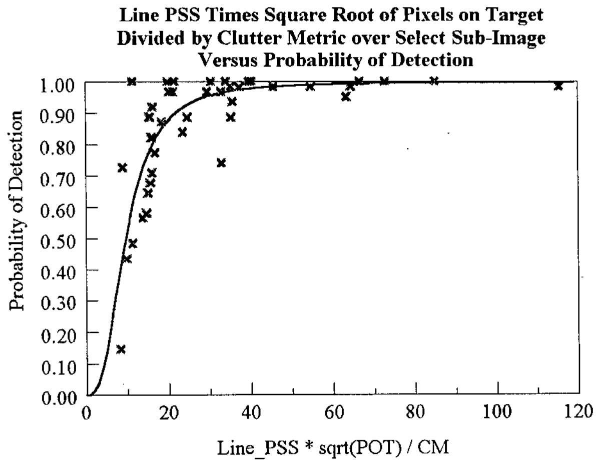

Figure 4a: PSS•√POT/CM2 vs. Probability of Detection

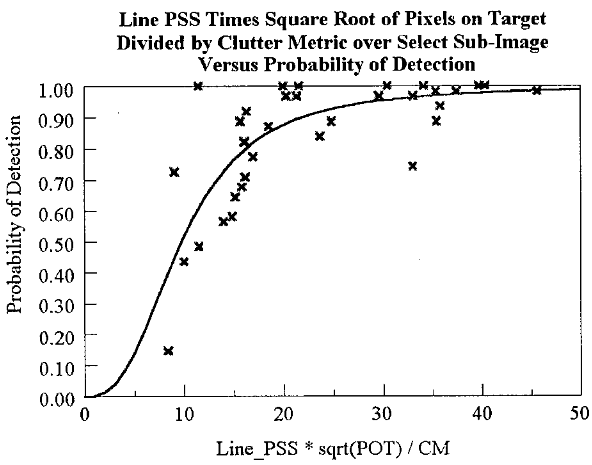

Figure 4b: PSS•√POT/CM2 vs. Probability of Detection enlargement of left side of Figure 4a.

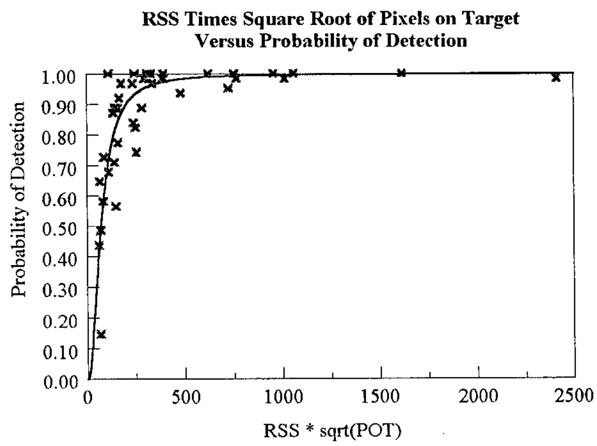

Figure 5a: RSS•√POT vs. Probability of Detection

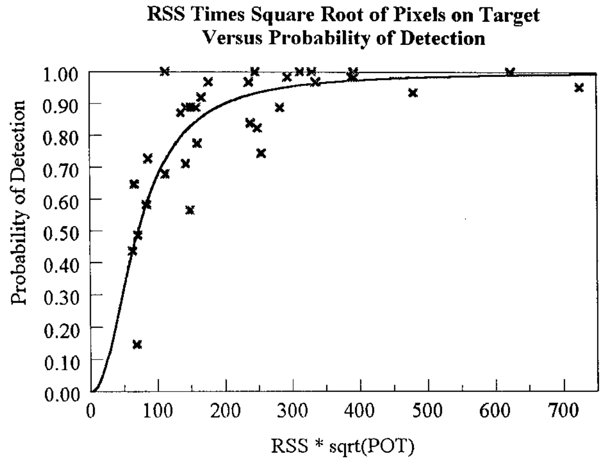

Figure 5b: RSS•√POT vs. Probability of Detection enlargement of left side of Figure 5a.

The outlier with probability of detection equal to 1.00 and least prediction metric in each case is image 6. This target casts a very conspicuous shadow that draws an observer's attention. Although the shadow was segmented with the target in the line PSS calculation, the shadow is probably a very strong cue to the observer.

5. CONCLUSION

The metrics used in this paper perform well enough against the Search-2 data to warrant further investigation in predicting visible probability of detection. The proposed contrast-to-clutter methodology has the promise of being somewhat self-calibrating. The studied metrics do not include color. Also, the metrics in the modeling presented do not include shape, although shape is also obviously important in target detection. Attempts to use a subjective shape measurement (Ref. 2) were unsuccessful with this data set. This may be due to the rather limited data set or the correlation between size and observer recognition of shape. The roles of contrast, size, shape and color are complicated by correlation between these factors. Additional research is needed that realizes this correlation.

6. REFERENCES

1. D'Agostino, J., Lawson, W., Wilson, D., "Concepts for search and detection model improvements", *Proceedings of SPIE*, Vol. 3063, pp.,14-22, 1997. Also available on the CD-ROM *Selected SPIE Papers on CD-ROM: Infared Technology 1988-1998,* Vol. 1, 1998.
2. D'Agostino, J. A., Wilson, D. L., Blecha, B. A., Biederman, I., Lawson, W. R., "The Role of Shape in Human Search & Detection", *21st Army Science Conference Proceedings* , 1998.
3. J. D. Silk, *Statistical Variance Analysis of Clutter Scenes and Applications to a Target Acquisition Test*, IDA Paper P-2950, November 1995.
4. Schmieder, D. E., Weathersby, M. R., "Detection Performance in Clutter with Variable Resolution", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 19, No. 4, p. 622, July 1983.